Forget those $%@*!?& commands: The Practical Supercomputing Toolkit

http://www.pstoolkit.org

Joe Werne, Michael Gourlay, Chris Meyer, Chris Bizon
Colorado Research Associates Division, NorthWest Research Associates, Inc.

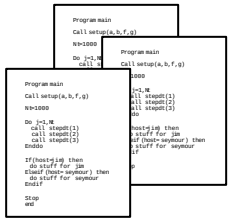
Aram Kevorkian (Chair, Metacomputing Working Group), Bill Asbury, Anthony DelSorbo (ASC), Virginia Bedford (ARSC), Bradford Blasing (AHPCRC), Rebecca Fahey, David Sanders (ERDC), Mitch Murphy (MHPCC), John Skinner (NAVO), Steve Thompson (ARL), Mark Dotson, Winfried Bernhard (WPAFB)

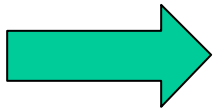
... and the users who have provided valuable feedback



The Problem

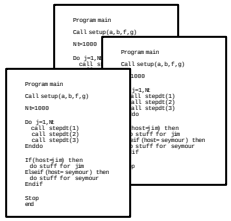
The Problem

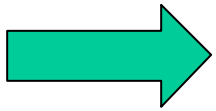




The Problem

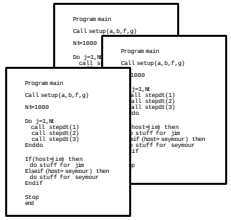
Code A

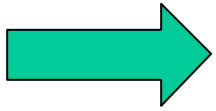




Cite A

Code B

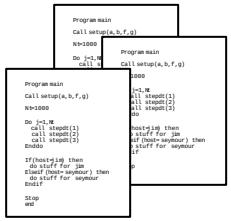


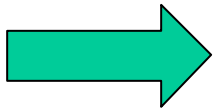


Cite B

The Problem

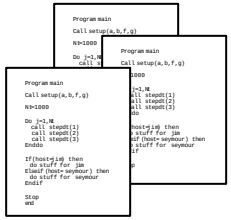
Code A

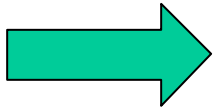




Cite A

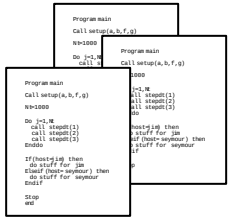
Code B

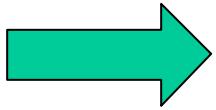




Cite B

Code C





Cite C

The Problem

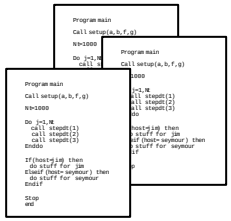
host	mkdir	put/get	migrate	queue
ERDC/jim	msfmkdir	msfput, msfget msfmput, msfmget	msfdmput, msfdmget	PBS
NAVO/seymour	remsh jules mkdir remsh vincent mkdir	rcp $f jules-hip0:$d rcp $f vincent-hip0:$d	remsh jules release remsh vincent release	NQS
ASC/hpc03-1	mkdir	/usr/bsd/rcp $f $(msas):$ARC/$d		PBS
AHPCRC/md	mkdir	cp	dmput, dmget	NQS
PSC/jaromir	far mkdir	far store far get		NQS
SDSC/golden	pftp mkdir hsi mkdir	pftp cd $d, put pftp cd $d, get		NQS
UCAR/ouray	****	mswrite, msread	-offline	NQS

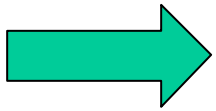
The Problem

- Unnecessary spin up time for new users,
- Dependence on specific sites/hardware,
- Reduced productivity,
- Increased project cost,
- Discouraged users,
- Duplicated effort.

The Problem

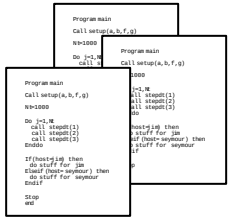
Code A

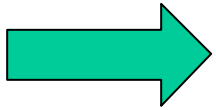




Cite A

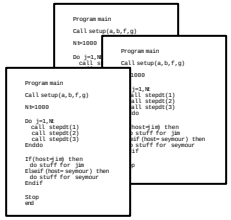
Code B

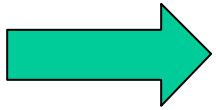




Cite B

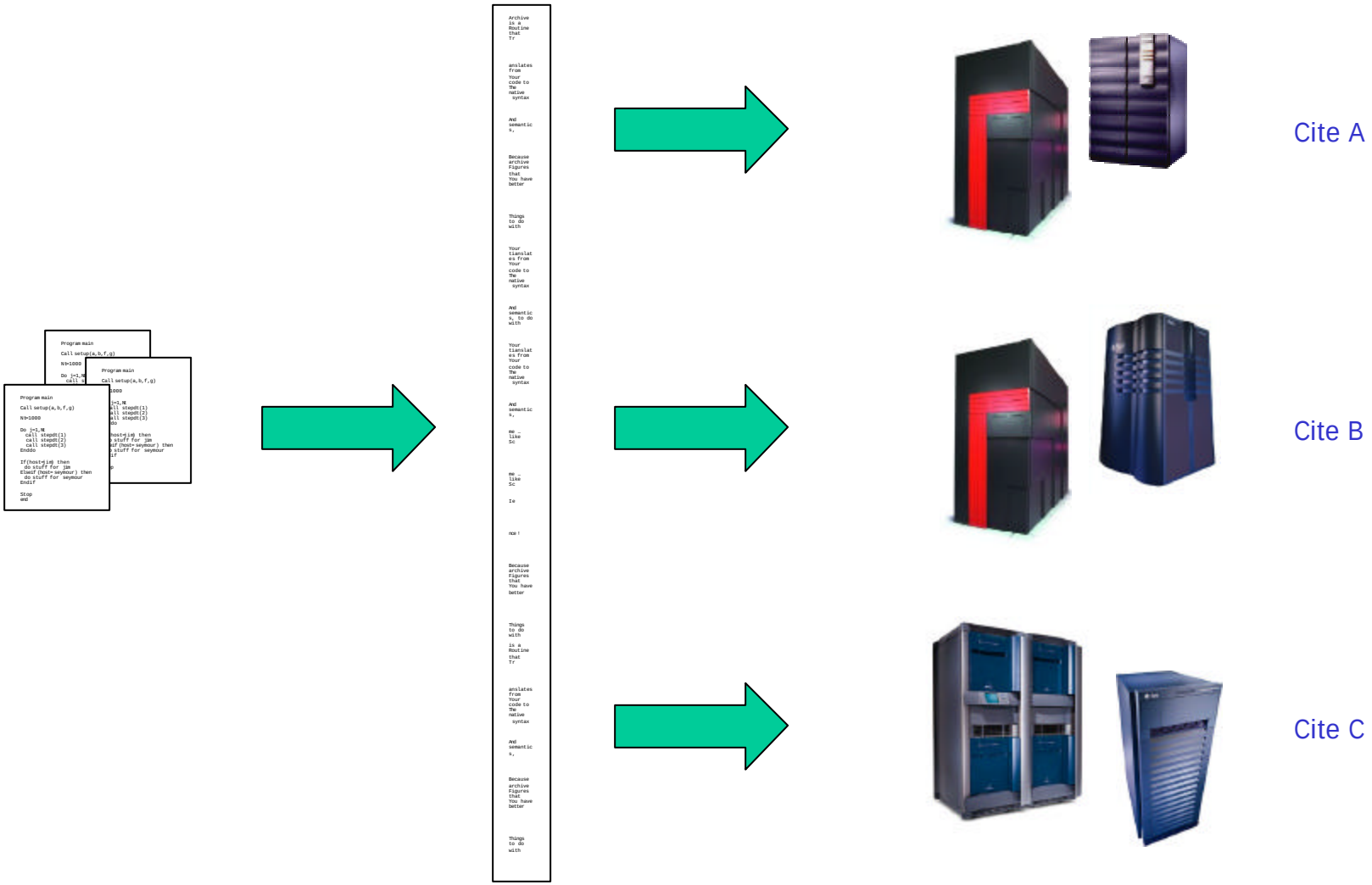
Code C





Cite C

The Solution



HPCMO & Metacomputing Working Group Initiative

Goals:

HPCMO & Metacomputing Working Group Initiative

Goals:

- Provide translation tools to all DoD users in the form of a uniform command-line interface.

HPCMO & Metacomputing Working Group Initiative

Goals:

- Provide translation tools to all DoD users in the form of a uniform command-line interface.
- Write nimble translation tools that are modular, maintainable, and reusable.

HPCMO & Metacomputing Working Group Initiative

Goals:

- Provide translation tools to all DoD users in the form of a uniform command-line interface.
- Write nimble translation tools that are modular, maintainable, and reusable.
- Shepherd implementation on all major U.S. super-computer centers – not just the DoD centers.

The Solution

How many translation tools do you need?

The Solution

How many translation tools do you need?

2

qprep, archive

The Solution

How many translation tools do you need?

2

qprep, archive

Unless you can think of more.

But can you?

The Solution

Tools for Uniform SuperComputing (TUSC): qprep

qprep:
simplifies and unifies job submission to queues

qprep is not a replacement for existing queuing systems, e.g., PBS, NQS, ...

rather, qprep is a translator between them.

The Solution

Tools for Uniform SuperComputing (TUSC): qprep

qprep is planned to work in two ways, both of which are intended to be familiar to users:

1. Command-line arguments:

```
qprep script –mpp_p 512 –mpp_t 12.45 …
```

2. Pseudo-comment directives in script preamble:

```
#PSTQ –l mpp_p 512
#PSTQ –l mpp_t 12.45
…
```

qprep will edit the script, translating the preamble to the native queuing system.

The Solution

Tools for Uniform SuperComputing (TUSC): qprep

The current plan is for implementation to be as simple as filling out a "translation table".

Prototype code for qprep is currently running on the T3E's at ERDC, NAVO, and AHPCRC, and on the Origin 2000 machines at ASC.

The Solution

Tools for Uniform SuperComputing (TUSC): archive

archive:
simplifies and unifies using data archives

archive provides a collection of commands to manipulate archival data, including put, get, ls, stat, mkdir, each of which accepts a list of command-line options.

Current plans include implementation of mv, rm, rmdir, chmod, chgrp, and chown as well.

The Solution

Tools for Uniform SuperComputing (TUSC): archive

archive is written in two layers:

- archive.pm is a Perl module containing subroutines that are site-independent.
- local.pm is a Perl module which contains subroutines that depend on the local system.

local.pm intentionally contains as little code as possible and ample comment statements to facilitate implementation at a new site.

The Solution

Tools for Uniform SuperComputing (TUSC): archive

Prototype code for archive is currently running on the T3E's at ERDC, NAVO, And AHPCRC, and on the Origin 2000 machines at ASC.

The following POC's for implementation, maintenance, and improvement of archive and qprep have been assembled: Steve Thompson (ARL), Bill Asbury, Anthony DelSorbo (ASC), Virginia Bedford (ARSC), Bradford Blasing (AHPCRC), Rebecca Fahey, David Sanders (ERDC), Mitch Murphy (MHPCC), John Skinner (NAVO), Mark Dotson, and Winfried Bernhard (WPAFB).

The Practical Supercomputing Toolkit

To facilitate implementation and distribution of the DoD uniform command-line interface, including archive and qprep ...

the HPCMO has funded the development of the Practical Supercomputing Toolkit (PST).

The Practical Supercomputing Toolkit

We envision PST will be:

The Practical Supercomputing Toolkit

We envision PST will be:

- A collection of solutions to common logistical problems encountered by users at supercomputing centers, allowing users to operate at a higher level of abstraction.

The Practical Supercomputing Toolkit

We envision PST will be:

- A collection of solutions to common logistical problems encountered by users at supercomputing centers, allowing users to operate at a higher level of abstraction.

- Open source, allowing users to contribute their own solutions to problems for the good of the user community, and in turn benefit from the solutions of others.

The Practical Supercomputing Toolkit

We envision PST will be:

- A collection of solutions to common logistical problems encountered by users at supercomputing centers, allowing users to operate at a higher level of abstraction.
- Open source, allowing users to contribute their own solutions to problems for the good of the user community, and in turn benefit from the solutions of others.
- Freely available from www.pstoolkit.org.

The Practical Supercomputing Toolkit

We envision PST will be:

- A collection of solutions to common logistical problems encountered by users at supercomputing centers, allowing users to operate at a higher level of abstraction.
- Open source, allowing users to contribute their own solutions to problems for the good of the user community, and in turn benefit from the solutions of others.
- Freely available from www.pstoolkit.org.
- Written primarily in Perl so as to benefit from Perl's advanced text processing capability, vast user support, and its Plain Old Documentation (POD) system. With POD, source code comments can be easily used to generate documentation as man pages, HTML, LaTeX, or Framemaker.

The Practical Supercomputing Toolkit

We envision PST will be:

- A collection of solutions to common logistical problems encountered by users at supercomputing centers, allowing users to operate at a higher level of abstraction.
- Open source, allowing users to contribute their own solutions to problems for the good of the user community, and in turn benefit from the solutions of others.
- Freely available from www.pstoolkit.org.
- Written primarily in Perl so as to benefit from Perl's advanced text processing capability, vast user support, and its Plain Old Documentation (POD) system. With POD, source code comments can be easily used to generate documentation as man pages, HTML, LaTeX, or Framemaker.
- Soliciting users for input on requirements for tools.

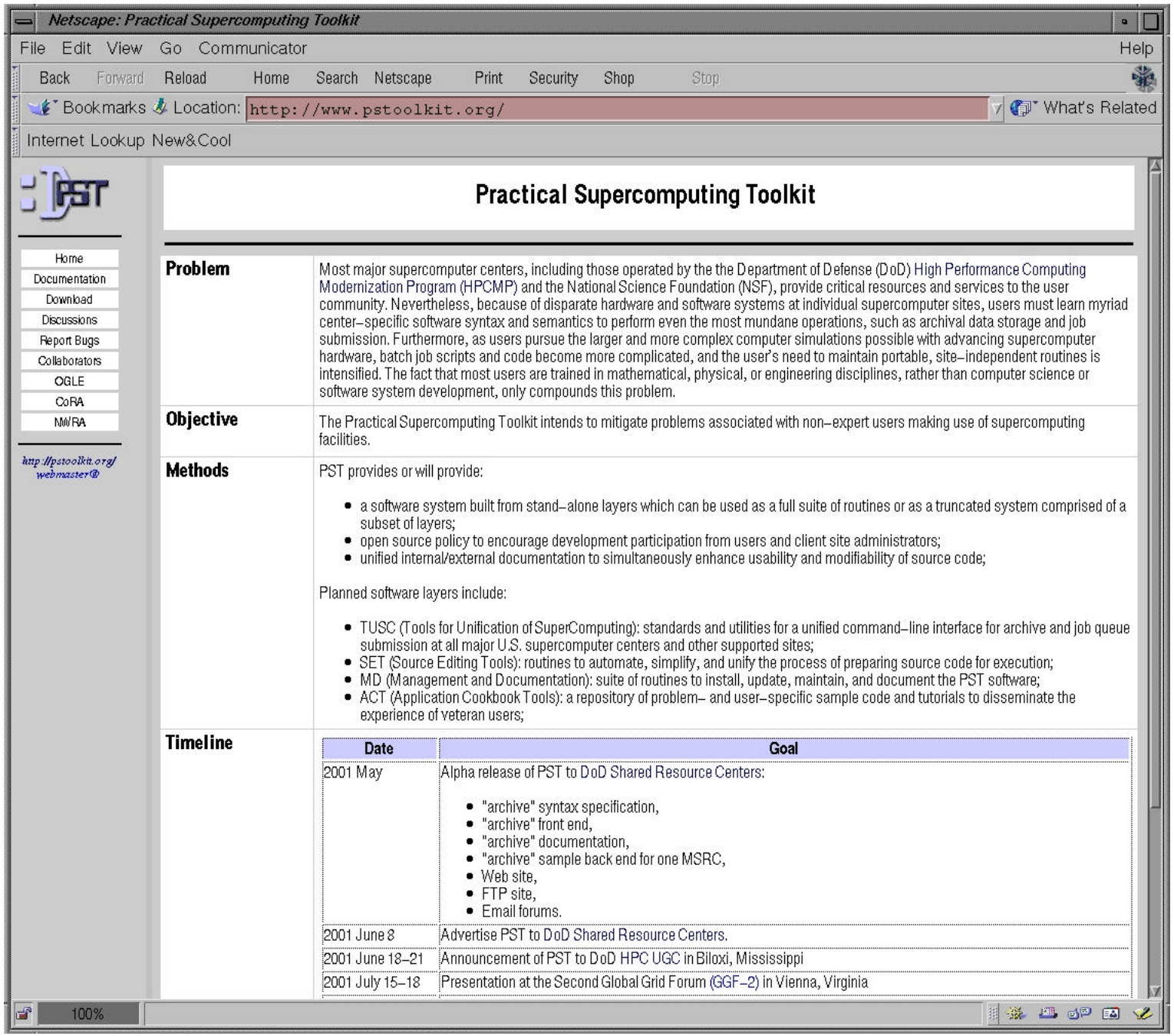
The PST Layers

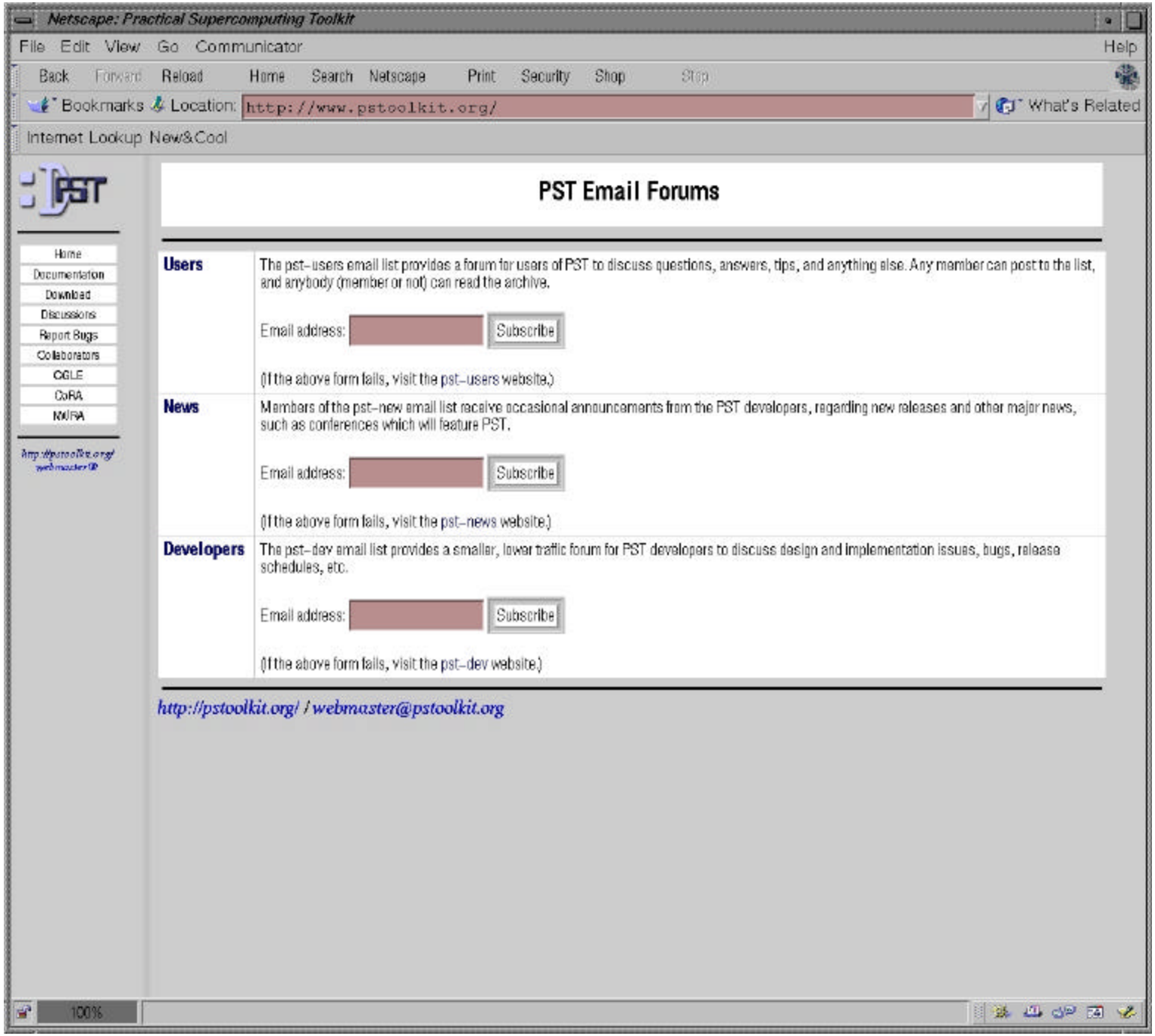
TUSC (Tools for Uniform Supercomputing)
Translation tools (including qprep and archive) that hide site-specific details.

SET (Source Editing Tools)
Tools that simplify the process of preparing source code for execution.

MD (Maintenance and Documentation)
Routines to install, update, and document PST.

ACT (Application Cookbook Tools)
A repository of problem-specific samples and tutorials built with TUSC and SET. A collection of contributions by veteran users that gathers and disseminates experience.


Netscape: Practical Supercomputing Toolkit
File Edit View Go Communicator Help
Back Forward Reload Home Search Netscape Print Security Shop Stop
Bookmarks Location: http://www.pstoolkit.org/ What's Related
Internet Lookup New&Cool
Home
Documentation
Download
Discussions
Report Bugs
Collaborators
OGLE
CoRA
NWRA
http://pstoolkit.org/
webmaster@
Practical Supercomputing Toolkit
Problem
Most major supercomputer centers, including those operated by the the Department of Defense (DoD) High Performance Computing Modernization Program (HPCMP) and the National Science Foundation (NSF), provide critical resources and services to the user community. Nevertheless, because of disparate hardware and software systems at individual supercomputer sites, users must learn myriad center–specific software syntax and semantics to perform even the most mundane operations, such as archival data storage and job submission. Furthermore, as users pursue the larger and more complex computer simulations possible with advancing supercomputer hardware, batch job scripts and code become more complicated, and the user's need to maintain portable, site–independent routines is intensified. The fact that most users are trained in mathematical, physical, or engineering disciplines, rather than computer science or software system development, only compounds this problem.
Objective
The Practical Supercomputing Toolkit intends to mitigate problems associated with non–expert users making use of supercomputing facilities.
Methods
PST provides or will provide:
a software system built from stand–alone layers which can be used as a full suite of routines or as a truncated system comprised of a subset of layers;
open source policy to encourage development participation from users and client site administrators;
unified internal/external documentation to simultaneously enhance usability and modifiability of source code;
Planned software layers include:
TUSC (Tools for Unification of SuperComputing): standards and utilities for a unified command–line interface for archive and job queue submission at all major U.S. supercomputer centers and other supported sites;
SET (Source Editing Tools): routines to automate, simplify, and unify the process of preparing source code for execution;
MD (Management and Documentation): suite of routines to install, update, maintain, and document the PST software;
ACT (Application Cookbook Tools): a repository of problem– and user–specific sample code and tutorials to disseminate the experience of veteran users;
Timeline
Date
Goal
2001 May
Alpha release of PST to DoD Shared Resource Centers:
"archive" syntax specification,
"archive" front end,
"archive" documentation,
"archive" sample back end for one MSRC,
Web site,
FTP site,
Email forums.
2001 June 8
Advertise PST to DoD Shared Resource Centers.
2001 June 18–21
Announcement of PST to DoD HPC UGC in Biloxi, Mississippi
2001 July 15–18
Presentation at the Second Global Grid Forum (GGF–2) in Vienna, Virginia
100%


Netscape: Practical Supercomputing Toolkit
File Edit View Go Communicator Help
Back Forward Reload Home Search Netscape Print Security Shop Stop
Bookmarks Location: http://www.pstoolkit.org/ What's Related
Internet Lookup New&Cool
Home
Documentation
Download
Discussions
Report Bugs
Collaborators
OGLE
CoRA
NWRA
PST Email Forums
Users
The pst-users email list provides a forum for users of PST to discuss questions, answers, tips, and anything else. Any member can post to the list, and anybody (member or not) can read the archive.
Email address: Subscribe
(If the above form fails, visit the pst-users website.)
News
Members of the pst-new email list receive occasional announcements from the PST developers, regarding new releases and other major news, such as conferences which will feature PST.
Email address: Subscribe
(If the above form fails, visit the pst-news website.)
Developers
The pst-dev email list provides a smaller, lower traffic forum for PST developers to discuss design and implementation issues, bugs, release schedules, etc.
Email address: Subscribe
(If the above form fails, visit the pst-dev website.)
http://pstoolkit.org/ / webmaster@pstoolkit.org
100%

The Future of PST

Over the next 18 months, tools will be implemented, released, and refined based on user input, and the number of centers that support PST will increase, aided by the open source philosophy.

July 15, 2001	Initial release of qprep man page.
December, 2001	qprep version 1.0 release.
May, 2002	Initial release of MD and SET layers.
October, 2002	Full release of TUSC, MD, and SET layers.

ACT layer codes will be released as contributed by users.

You can impact the development of PST

You can impact the development of PST

- The evolution of PST relies on users for feedback on tool definition and implementation, as well as code submissions from user/developers.

You can impact the development of PST

- The evolution of PST relies on users for feedback on tool definition and implementation, as well as code submissions from user/developers.
- Review the on-line archive man page and offer feedback.

You can impact the development of PST

- The evolution of PST relies on users for feedback on tool definition and implementation, as well as code submissions from user/developers.
- Review the on-line archive man page and offer feedback.
- Look for qprep documentation for review on July 15, 2001.

You can impact the development of PST

- The evolution of PST relies on users for feedback on tool definition and implementation, as well as code submissions from user/developers.
- Review the on-line archive man page and offer feedback.
- Look for qprep documentation for review on July 15, 2001.
- Sign up for one of pstoolkit's three mailing lists: pst-news for release announcements; pst-users for user discussions, questions, answers, and tips; pst-dev for discussions of implementation, bugs, and development logistics.

You can impact the development of PST

- The evolution of PST relies on users for feedback on tool definition and implementation, as well as code submissions from user/developers.
- Review the on-line archive man page and offer feedback.
- Look for qprep documentation for review on July 15, 2001.
- Sign up for one of pstoolkit's three mailing lists: pst-news for release announcements; pst-users for user discussions, questions, answers, and tips; pst-dev for discussions of implementation, bugs, and development logistics.
- Early adopters can download PST from the website, try out archive, and provide feedback. Currently, archive is implemented for ERDC, AHPCRC, ASC, and NAVO.

You can impact the development of PST

- The evolution of PST relies on users for feedback on tool definition and implementation, as well as code submissions from user/developers.
- Review the on-line archive man page and offer feedback.
- Look for qprep documentation for review on July 15, 2001.
- Sign up for one of pstoolkit's three mailing lists: pst-news for release announcements; pst-users for user discussions, questions, answers, and tips; pst-dev for discussions of implementation, bugs, and development logistics.
- Early adopters can download PST from the website, try out archive, and provide feedback. Currently, archive is implemented for ERDC, AHPCRC, ASC, and NAVO.
- Because PST is open source, developers at other centers can contribute implementations of archive, which will then be added to the full distribution.

Acknowledgments

The DoD High Performance Computing Modernization Office provides the initial funding of the TUSC, SET, and MD layers of PST, with emphasis on unification across DoD Shared Resource Centers for job submission and data archiving processes. The DoD Metacomputing Working Group provides logistical coordination to encourage the participation of the DoD centers.